Genomics 97 (2011) 304–312

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Genomic evolution in domestic cattle: Ancestral haplotypes and healthy beef

Joseph F. Williamson [a,b], Edward J. Steele [a,b], Susan Lester [c], Oscar Kalai [a], John A. Millman [a], Lindsay Wolrige [a], Dominic Bayard [d], Craig McLure [a,b,e,g], Roger L. Dawkins [a,b,f,*]

[a] C.Y. O'Connor ERADE Village Foundation, Canning Vale, Western Australia, Australia
[b] Division of Health Sciences, Murdoch University, Murdoch, Western Australia, Australia
[c] Rheumatology Department, The Queen Elizabeth Hospital, Woodville, South Australia, Australia
[d] Global Reproduction Solutions, Goorambat, Victoria, Australia
[e] Genetic Technologies Ltd, Fitzroy, Victoria, Australia
[f] Faculty of Medicine and Dentistry, University of Western Australia, Nedlands, Western Australia, Australia
[g] Department of Veterinary Sciences, University of Melbourne, Victoria 3000, Australia

ARTICLE INFO

Article history:
Received 21 December 2010
Accepted 9 February 2011
Available online 19 February 2011

Received by J. Quakenbush

Keywords:
Bovine genomics
SREBF1–GH polymorphisms
Ancestral haplotypes
Meat quality
Simmental, Angus, Wagyu cattle
Genomics of cattle breeds

ABSTRACT

We have identified numerous Ancestral Haplotypes encoding a 14-Mb region of Bota C19. Three are frequent in Simmental, Angus and Wagyu and have been conserved since common progenitor populations. Others are more relevant to the differences between these 3 breeds including fat content and distribution in muscle. SREBF1 and Growth Hormone, which have been implicated in the production of healthy beef, are included within these haplotypes. However, we conclude that alleles at these 2 loci are less important than other sequences within the haplotypes. Identification of breeds and hybrids is improved by using haplotypes rather than individual alleles.

1. Introduction

Domestic cattle, *Bos taurus* (Bota), have coevolved with humans for some thousands of years during which intense selection has been practised. Many distinct breeds are maintained but for different purposes in diverse environments. There must have been some "pure" geographically isolated breeds but few, if any, of these still exist. Reconstructing the evolution of current cattle will reveal fundamental aspects of genomics including mechanisms relating to survival and change. Issues such as the conservation of polymorphism, gene interaction and hybrid vigor are of practical importance and can be approached by defining the evolution of founder or progenitor populations, successful descendants (or "breeds") and extant mixes now available for study.

Understanding the genetic control of beef production has been addressed in many previous studies. Most have sought to identify traits which might assist the cattle producer to improve efficiency but some have focused on consumer preferences including variables such as taste, tenderness, intramuscular fat or "marbling". The 'holy grail' is the identification of genes which result in "healthy" beef in terms of lower melting point fat content, type and distribution.

Undoubtedly many of these characteristics are under complex polygenic control. The challenges are compounded by the diversity of cattle breeds which have been selected by phenotype for hundreds of generations. Desirable features have been achieved by different routes and on different ancestral backgrounds.

Some genetic tests have been offered commercially, but, few, if any, are of practical value. Independent studies have failed to support the claims of the commercializing company [1]. Such failure appears to be due to the lack of strong candidate genes, the complexity of the genetics and a dependence on statistical analysis of association and linkage. It is hardly surprising that promising results in one breed do not relate to another.

There is vastly more experience in identifying the genetic control of health and disease in *Homo sapiens* (Hosa). Successful studies were readily possible in the case of monogenic disorders but it has been far more difficult in the case of polygenic traits. In spite of huge expenditure, the yield from Genome Wide Analyses (GWA) of Single Nucleotide Polymorphisms (SNPs) remains disappointing.

In considering a strategy to apply to cattle, we have taken advantage of the results which have followed extensive multicenter analyses of the Hosa Major Histocompatibility Complex (MHC). Some of the conclusions are reviewed in detail elsewhere [2–7].

* Corresponding author at: C.Y. O'Connor ERADE Village Foundation, Canning Vale, Western Australia, Australia.
E-mail address: rldawkins@cyo.edu.au (R.L. Dawkins).

0888-7543/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2011.02.006

Undoubtedly, however, the key is to understand that the alternative megabase sequences are both polymorphic and conserved. A limited number of polymorphisms can define the entire haplospecific sequence which has been inherited faithfully from remote ancestors. Such ancestral haplotypes encode multiple haplospecific coding and 'non-coding' sequences which together control traits and diseases. Current phenotypes reflect the mix of ancestral haplotypes at all polymorphic frozen blocks. Alper et al. [7] and Smith et al. [2] have demonstrated the benefit of using these observed haplotypes rather than the GWA and SNP approach.

So as to focus on ancestral haplotypes relevant to the production of "healthy" beef we needed to study a genomic region containing one or more polymorphic candidate genes. We selected a section of Bota C19 containing Sterol Regulatory Element Binding Transcription Factor 1 (SREBF1) and Growth Hormone (GH1) since both have been promoted as having some form of commercial utility (see Fig. 1).

SREBF1 is a transcription factor which regulates gene expression levels of Stearoyl-CoA Desaturase (SCD) and other genes relevant to lipid and fatty acid metabolism in tissue [8]. The content of unsaturated fatty acid in beef is an important factor in taste and texture [8,9] and the degree of marbling. A recent study revealed that high SCD activities were correlated positively with beef marbling score, amount of monounsaturated fatty acid and conjugated linoleic acid, but negatively with the amount of saturated fatty acid [10]. SREBF1 may be particularly important [9] but we conclude that undiscovered closely linked genes may be more so. Japanese Wagyu with highly marbled beef have a high frequency of the short allele (S) of SREBF1 unlike Angus which are also known for their marbling, albeit different in content and distribution. We predicted that breed-specific haplotypes, unlike SREBF1 alleles alone, would define helpful multilocus markers and ultimately identify the many sequences which together control the character of beef. GH typing is in demand, at least partly, as an independent guide to performance in lot feeding.

In this study we used in-house methods of primer design to permit ancestral haplotyping [4,5,11] of cattle breeds with very different degrees of marbling, *inter alia*.

2. Results

2.1. Aims and strategy

With the intention of comparing and contrasting different but highly selected breeds and of relating genomic polymorphism to "healthy" beef, we took advantage of our relatively closed, relatively pure breeds of European Simmental, Japanese Wagyu and British Angus cattle. In the case of Simmental, pedigrees have been kept for 40 years. Selection criteria include fertility, polling, temperament, size and muscling without fat. DNA has been collected for 20 years. The Angus have been under intense phenotypic selection for many of the same traits but particularly the ability to fatten and marble on grass. The Wagyu are derived from semen and embryos imported from Japan within the last 10 years and are known for an ability to marble after many months of lot feeding. The descendants are thought to be representative of Japanese Blacks, which, it is believed, have been selected for low melting temperature, 'healthy' intramuscular fat. In essence, Simmentals are large, strong, hardy and lean, Angus are short, fat and tasty, by comparison, and Wagyu are small, fragile, slowly maturing and preferred in terms of their ability to deposit low melting temperature fat within muscle.

The strategy assumes that these three herds must show genomic differences to reflect their selection and current phenotypes. However, many of these differences will reflect the background and ancestry whereas we intend to identify those which control the preferred phenotypes. It follows that earlier results have the potential to obfuscate and to lead to premature commercialization, as has happened with monogenic and SNP approaches.

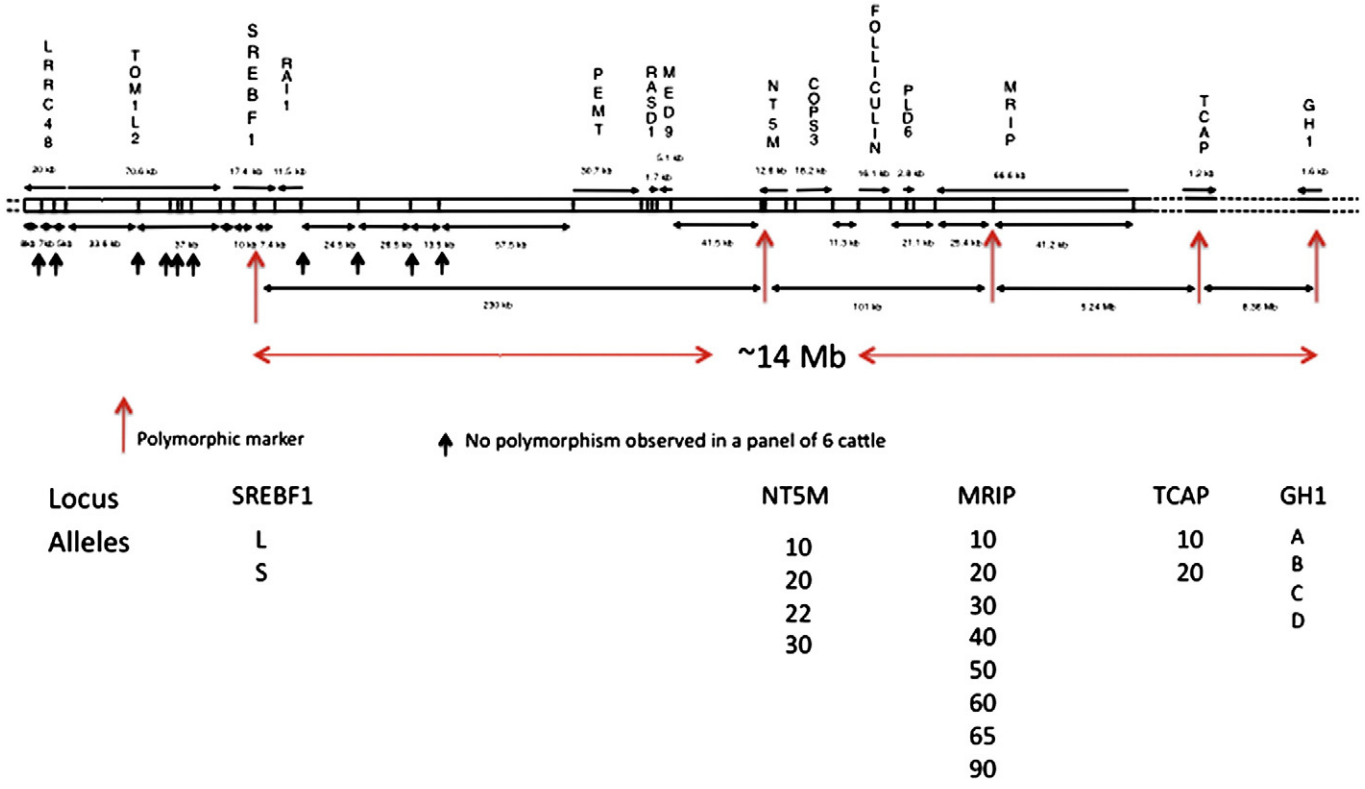


Fig. 1. The genomic map around SREBF1 to GH1 on C19. The red vertical arrows show the positions of the five markers used in this study. The black vertical arrows show the positions of other primers which were not polymorphic when tested on a screening panel of 6 animals. Horizontal arrows shown above the chromosome indicate the orientation of transcription. SREBF1—sterol regulatory element binding transcription factor 1; NT5M—nucleotidase dephosphorylates uracil and thymine; MRIP—myosin phosphatase-Rho interacting protein; TCAP—titin-cap; GH—growth hormone.

Undoubtedly, the genomic differences we seek are more likely to be found by using the *observed haplotype* approach which has proven successful in humans [7]. It is quite clear that the genome is arranged into polymorphic frozen blocks encoding multiple ancestral haplotypes which are inherited en bloc [12,13]. It is also clear that the complex polygenic traits we intend to define are dependent on *cis*, *trans* and epistatic interactions.

So as to reduce the risk of confusion, we have adopted a matrix which provides criteria for interpretation of genomic differences. For example, Angus and Wagyu should share some sequences relating to fat deposition and should share differences with Simmental. Wagyu and Angus should be different in the sequences which determine distribution and melting temperature of fat and these sequences should be different again in Simmental.

The strategy also requires the identification of haplospecific combinations of markers which identify long sequences without knowledge as to how each nucleotide functions or varies. Even if all subjects were sequenced in entirety, it is not obvious how the analysis could proceed expeditiously without comparing the phenotypic associations of ancestral haplotypes of megabase length.

To these ends, four steps are used.

First, identify homozygous haplotypes sharing paternal and maternal copies of the same markers and therefore intervening sequence over extended distances as shown by us and other groups [2,4,7,14].

Second, confirm these putative haplotypes by showing faithful inheritance over at least 3 generations, thereby excluding recent recombinants.

Third, after defining the more common haplotypes, it is then possible to assign other less common haplotypes in heterozygotes.

Last, determine which of these haplotypes are ancestral in the sense that they exist in at least several subjects and families which can only be related by remote descent thereby implying conservation from distant ancestors.

2.2. Haplotyping of the genomic region

The genomic map around SREBF1 is shown in Fig. 1. It can be seen that there are several loci which could be relevant. For example, MRIP is myosin phosphatase Rho interacting protein and TCAP is titin-cap (telethonin), which are relevant to muscle function. The GH locus is some 14 Mb from SREBF1 and has been assumed to segregate independently.

So as to test the extent of haplotypes covering the region, we designed PCR primers around NT5M, MRIP and TCAP as well as SREBF1 and GH. For convenience the number of alleles at each of the 5 loci is listed.

Representative gel images for amplified products are illustrated in Fig. 2 and tabulated results for all tested cattle are shown in Table 2. The band intensities are given as 1 to 9 such that 1 represents negative (not shown here), 2 represents equivocal and 3 to 9 are weak to maximal positive.

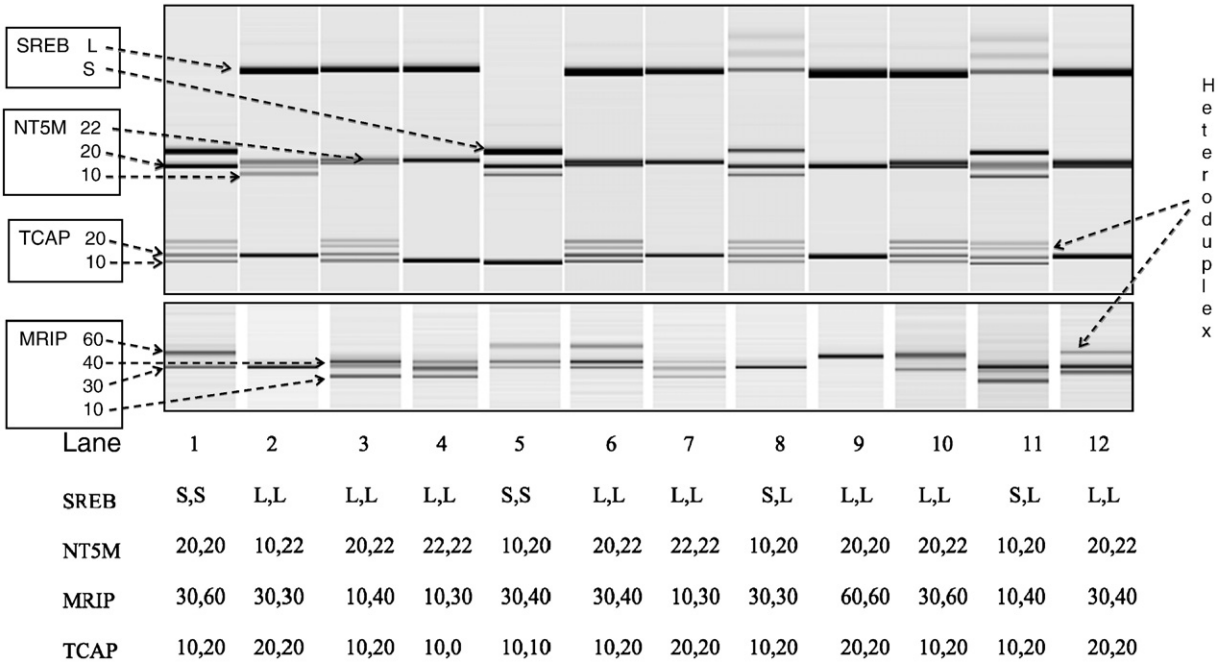
Table 1
Primer sequences for C19 haplotyping.

SREBP1 F[a]	5′ CCA CAA CGC CAT CGA GAA ACG CTA C 3′
SREBP1 R[a]	5′ GGC CTT CCC TGA CCA CCC AAC TTA G 3′
NT5M F	5′ GGA AGG CCA GTT ACA TGG CA 3′
NT5M R	5′ TGT GAT TTG GCC TTG GTT GTG 3′
MRIP F	5′ AGG GGT GCT GAG YCT ACA GG 3′
MRIP R	5′ CTC CAG GAG GCA GGA GAA G 3′
TCAP F	5′ AGT ACC AGC TGC CCT ACC A 3′
TCAP R	5′ CTG AGA CAT GGA GCG AGC CA 3′
GH F[b]	5′ TCT ATG AGA AGC TGA AGG ACC TGG AGG AA 3′
GH R1	5′ GGG GGG TGC CAT CTT CCA G 3′
GH R2[b]	5′ GGG GGG TGC CAT CTT CCA C 3′
GH R3[b]	5′ ATG ACC CTC AGG TAC GTC TCC G 3′
GH R4	5′ CAT GAC CCT CAG GTA CGT CTC CA 3′

[a] From Ref. [8].
[b] From Ref. [22].

2.3. Allele frequencies

As shown in Table 3, there are differences between European, Japanese and British breeds. Some of these are encouraging. For example, the three breeds are very different in allele frequencies at SREBF1 S, NT5M 22, MRIP 60 and GH C, which could lead to the conclusion that these alleles and statistically derived haplotypes



Lane	1	2	3	4	5	6	7	8	9	10	11	12
SREB	S,S	L,L	L,L	L,L	S,S	L,L	L,L	S,L	L,L	L,L	S,L	L,L
NT5M	20,20	10,22	20,22	22,22	10,20	20,22	22,22	10,20	20,20	20,22	10,20	20,22
MRIP	30,60	30,30	10,40	10,30	30,40	30,40	10,30	30,30	60,60	30,60	10,40	30,40
TCAP	10,20	20,20	10,20	10,0	10,10	10,20	20,20	10,20	20,20	10,20	10,20	20,20

Fig. 2. Examples of PCR products detected using QIAxcel high resolution non-denaturing capillary gels. SREBF1, NT5M and TCAP primers are combined in a multiplex PCR. Interpretation for each lane is shown. Note only certain heterozygotes produce heteroduplexes.

Table 2
Tabulation of banding pattern as shown in Fig. 2. Each lane is a different sample. Allele nomenclature is derived from the relative product size at each locus. The numbers under each allele heading represent the band intensity from weak to maximal positive (3–9). Blanks represent negatives.

Lane	SREBF1		NT5M			MRIP						TCAP			
	S	L	10	20	22	10	30	40	60	80	90	10	20	30	40
1	9			7			5		6			5	5	4	4
2		9	5		5		6						7		
3		9		5	5	5		7				5	5	4	4
4		9			7	5	6	4				7			
5	9		4	6			4	5		4		7			
6		9		6	7		5	7		5		5	5	4	4
7		9			7	3	4	3					7		
8	5	5	4	6			6					5	5	4	4
9		9		8					6				7		
10		9		6	7		5		5			5	5	4	4
11	7	6	5	5		5		6				5	5	4	4
12		9		6	6		5	6		4			7		

(such as S,22,60,X,C) are responsible for the differences between the 3 breeds.

2.4. Haplotypes

In keeping with the approach used to define MHC ancestral haplotypes in humans, we adopted a stepwise approach. First, individual subjects were classified as homozygous or heterozygous at each locus. The presence of heteroduplexes confirmed heterozygosity in some combinations at the SREBF1, MRIP and TCAP loci.

Next, we examined segregation in nuclear families. The tree shown in Fig. 3 includes pure Simmentals crossed with Angus and Dexter. The L,20,40,20,A is present in 2 unrelated Simmental bulls (9242 and 9764), in unrelated cows (9913) and some of their descendents (0174, 0286, 0058). The Angus bull (ARNIE L;22;30,40;20;A) is heterozygous for MRIP indicating the presence of L,22,30,20,A and L,22,40,20,A. The latter is found in 0308 together with L,20,40,20,A contributed by 9764 via 0058. The Dexter bull (Raff Dex3, L,22,30,20,A; L,20,40,20,A) has one ancestral haplotype found in Angus and another in Simmental suggesting that these 3 breeds have some common ancestry. Both segregate faithfully as illustrated by 0446 and 0411.

Finally, by following the same procedure in “purebreds” and their crosses, it is possible to list recurring haplotypes as in Table 4. Three (L,20,30,20,A and L,20,40,20,A and L,20,30,10,A) are present in more than 5% of Simmental, Angus and Wagyu, suggesting that these are shared by very remote ancestry from cattle progenitors (Table 4). Other haplotypes are largely breed specific (e.g. S,10,60,10,A and S,10,60,10,B in Wagyu) suggesting that these are shared by more recent ancestry since formation of the breed.

Some of the haplotypes listed in Table 4 are closely related and appear to have been generated by an ancient recombination as for example between TCAP and GH in the two Wagyu haplotypes sharing S,10,60,10 but differing in A and B at GH. There are other examples of apparent historic recombination such as in the frequent haplotypes sharing L,20,40,20 or L,20,30,20 but differing at GH. L,22,40,20,A is almost Angus specific but L,22,40,20 with B or C is Wagyu specific suggesting that recombination between TCAP and GH or mutation of GH A to B and C occurred after the separation of these two breeds.

2.5. Linkage disequilibrium

Common haplotypes can sometimes be identified by examining linkage disequilibrium for *cis* combinations which occur more frequently than expected from individual frequencies. However, this approach is less efficient than the direct method used above especially when:

more than 2 loci are required to define a haplotype,

rare haplotypes are critical,

populations are heterogeneous or ill defined, and

shared alleles occur at different relative frequencies.

As shown in Table 4, the most common or specific haplotypes (see Fig. 4) are suggested by the relatively high pair values (Table 5), reflecting L,20,30/40 and S,10,60 and 22,30/40, 20,A where “/” represents either/or and “X” unknown. However, as predicted from experience in humans, many uncommon haplotypes demonstrated by segregation are not revealed by linkage disequilibrium analysis and some expected, based on allele frequencies, such as X, 20,60,10, X, have not been seen. Ancestral haplotypes must be observed rather than predicted statistically.

Interestingly, the S,22,60,X,C haplotype suspected from allele differences (see above) does not occur in this data set of nearly 600 haplotypes. In fact, S and 22 and also 22 and 60 are in repulsion. The same allele can be carried on different haplotype sequences which are identified pragmatically by observing inheritance of specific combinations of alleles at linked loci e.g. S, non 22, 60. Genomic sequences are not revealed by allele frequencies and linkage disequilibrium. GWA studies can be very misleading and have also led to a distorted view of genome structure [2,7].

Table 3
Allele frequencies (%) for each locus in three cattle breeds.

Breed	SREBF1		NT5M			MRIP				TCAP		GH			
	S	L	10	20	22	10	30	40	60	10	20	A	B	C	n
Simmental	7.6	92.4	10.4	88.4	1.2	0	43.2	50.1	6.6	20.3	79.9	76.3	23.7	0	502
Wagyu	25.8	74.2	21	58	21	6.5	35.5	37.1	21	48.4	51.6	50	32.3	17.7	62
Angus	0	100	0	71.9	28.1	0	43.8	56.3	0	21.9	78.1	87.5	12.5	0	32

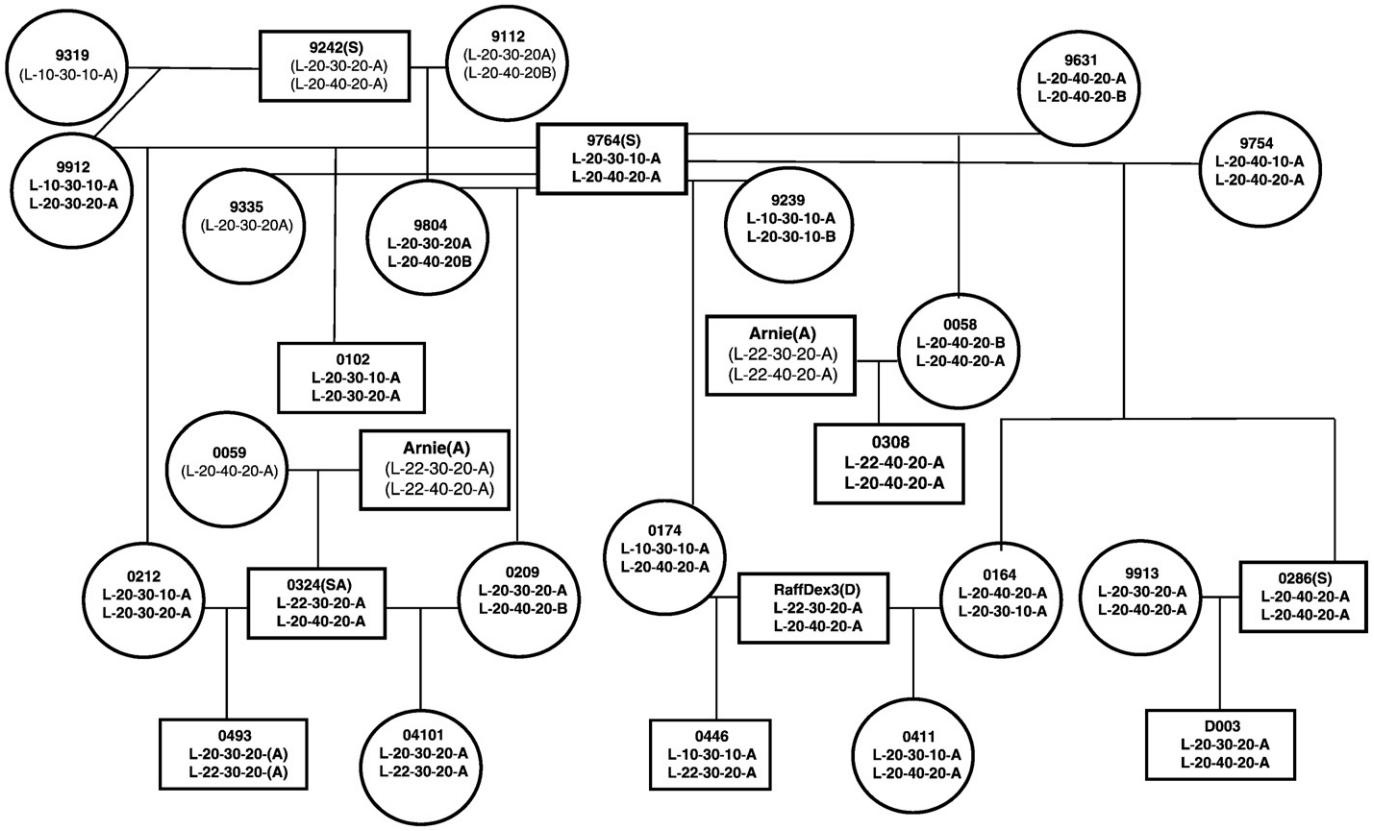


Fig. 3. Mendelian segregation in multigeneration families, including Simmental, Angus and Dexter cattle. The number above the haplotypes represents the cattle ID. Haplotypes/alleles in brackets have not been tested and hence are assumed. Breed of sires: (S)—Simmental, (A)—Angus, (SA)—Simmental-Angus, (D)—Dexter.

Table 4
Combined haplotype frequencies observed in 298 pure bred (Simmental $n=251$, Wagyu $n=31$, Aberdeen Angus $n=16$) where the haplotypes were shown by family segregation or could be unequivocally assigned (haplotypes observed where $n\geq 2$ (28 of 41)).

SREBF1	NT5M	MRIP	TCAP	GH	Freq %
L	20	40	20	A	25.8
L	20	30	20	A	21.3
L	20	30	10	A	6.9
L	20	40	20	B	8.7
L	20	30	20	B	7.0
L	20	40	10	A	6.2
L	10	30	10	A	2.9
L	20	60	20	A	2.2
L	22	40	20	A	2.0
L	20	30	10	B	1.7
S	10	60	20	A	1.7
S	10	60	10	B	1.5
S	10	60	20	B	1.2
S	10	30	20	A	1.0
S	20	40	20	B	0.8
L	22	40	20	C	0.8
L	10	40	10	A	0.8
L	10	40	10	B	0.7
L	20	10	10	A	0.5
L	20	40	10	B	0.5
L	22	40	20	B	0.5
S	10	60	10	A	0.5
S	20	40	20	A	0.5
S	20	30	20	A	0.5
S	20	30	20	B	0.3
L	22	40	10	A	0.3
L	22	40	10	B	0.3
L	20	40	20	C	0.3

2.6. Haplotype mapping

Recombinant haplotypes have proven useful in mapping functions and disease in humans. In the present study, the intention is to identify the regions which contribute to the breed characteristics of beef. Promising illustrative candidates include MRIP 60 to TCAP 10 and MRIP 40 to GH C for features specific to Wagyu and NT5M 22 to TCAP 20 for those shared by Wagyu and Angus but not Simmental. In this way, it is possible to build a matrix to assist the investigation of genomic sequences for their role in determining phenotypic similarities and differences.

2.7. Breed markers

Irrespective of their biological significance, observed haplotypes can be used to confirm or exclude certain breeds. Based on the present data alone, and presuming the aim is to distinguish between Wagyu and Angus, as is often required in practice, S,10,60 has a high positive predictive value (PPV) (see Fig. 5) for Wagyu and, more importantly, excludes Angus. In this example, the individual alleles would also be useful at least with respect to Wagyu versus Angus. However, utility cannot be determined by allele frequencies alone. For example TCAP 10 and GH B alleles are frequent in all three breeds but the 10, B together as a haplotype (or genomic segment) is almost restricted to Wagyu with a PPV over 90%.

In terms of certification of Angus, the L,22,40,20,A haplotype has a PPV of approximately 90% which is far higher than for each individual allele alone (see Table 3).

Although many haplotypes are not helpful in distinguishing between Wagyu and Simmental, including those common to all

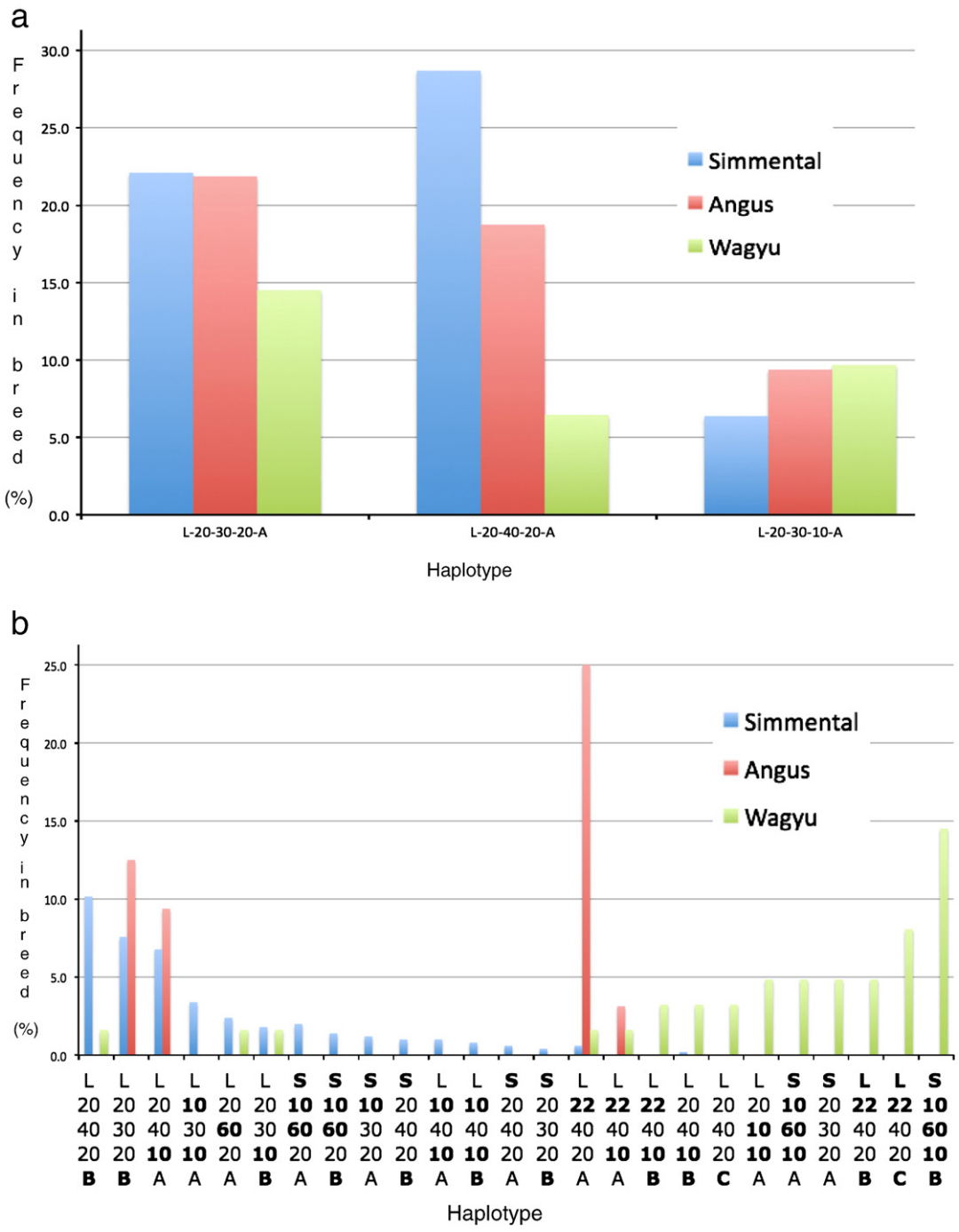


Fig. 4. Frequencies of Ancestral Haplotypes (AH) in Simmental, Aberdeen Angus and Wagyu cattle. (a) Haplotypes with frequencies ≥5% in all three breeds. (b) Haplotype frequency distribution in the three breeds where the haplotype was observed in 2 or more unrelated animals. Figures in bold represent emphasis of rarer alleles of interest.

breeds, some are potentially valuable depending upon the question and the operational definition of a breed.

2.8. Mixing breeds

An unexpected finding is the presence of S and S,10,60 in Simmental cattle. Pedigree analysis reveals that these markers are largely traceable to 2 elite bulls which are no longer available for study. From Table 4, it would appear that a S,10,60,20,A/B haplotype, which is similar but not identical to several Wagyu haplotypes, must have been introduced in recent times but long enough ago to permit recombination events which have resulted in fragmentation and the creation of S,10,X,X,X and S,X,X,X,X on a typical Simmental background. Such recombinants are invaluable reagents for gene discovery [15,16].

So as to identify further mixing and ultimately useful recombinants, we typed a panel including African, Asian and other European breeds. No new alleles were found with the exception of MRIP 90 in Zebu and Brahman and 50 and 65 in Brahman, further emphasizing the extreme polymorphism and the stability of this locus. Rare examples of S were present in Brahman and Jersey. Because family data were not available, we only assigned haplotypes in the case of homozygotes. All 3 of the haplotypes designated progenitor (L, 20,30,20 and L,20,40,20 and L, 20,30,10) are present in Brahman and probably Zebu in addition to the European breeds. The breed specificity of other haplotypes (see above) was confirmed.

Table 5
Two locus D′ in Simmental cattle ($n=234$): allele pairs showing significant positive associations.

SREBF1	NT5M	MRIP	TCAP	GH	D′	Freq
L	20				0.67	0.86
S	10				0.6	0.04
L		30			0.56	0.43
L		40			0.54	0.48
S		60			0.57	0.04
L			10		1.00	0.21
S			20		1.00	0.07
	10				0.44	0.03
	20				0.46	0.46
	22		20		1.00	0.01
	22			A	1.00	0.01
		60	20		1.00	0.06
			10	A	0.43	0.19
			20	B	0.43	0.02

2.9. Evidence for recent mutation

NT5M 22 occurs in Angus cattle and in breeds which are thought to have been infused with or derived from similar British stock. Wagyu are said to be Asian in origin but with contributions from Channel Island breeds such as Jersey and later contributions from Angus. It is possible that NT5M 22 arose within hundreds of generations. Supporting evidence is provided by the fact that 22 differs from 20 by only 2 basepairs. The GH alleles B and C may have arisen from A which could explain their distribution in Wagyu. The discovery of a new allele, GH D, in just one animal, is consistent with the fact that GH alleles are very similar as shown in Table 6a. MRIP 50, 65 and 90 in Brahman and Zebu could be recent mutations or, more likely, progenitor alleles which have been lost in the adaptation from tropical to moderate environments.

3. Discussion

Domestic cattle offer excellent opportunities to study evolution over the relatively short time period during which humans and their animals have migrated and adapted to changing environments while increasing their numbers exponentially. At issue is the extent of genomic change which has occurred contemporaneously. It would be helpful to know how current survivors compare with their ancestors prior to population expansion. There are two alternative scenarios. Firstly, under the popular change model, expansion has been associated with and dependent on change—sometimes referred to as "mutation"—which has improved fitness. Secondly, the conserved or predestined model says that change preceded and permitted expansion and migration.

Breed Specific Haplotypes

	SREBF1	NT5M	MRIP	TCAP	GH1	PPV%
Wagyu	S	10	60	x	x	85
	x	10	60	x	x	84
	x	20	10	10	x	98
	x	x	x	10	B	92
Simmental	x	0	30	10	x	93
	x	20	40	x	x	75
Angus	L	22	40	20	x	90
Zebu	L	x	90	x	x	70
Brahman	L	20	50	x	x	100

Fig. 5. Breed positive predictive values (PPV) of C19 haplotype.

Table 6a
Assignment of GH alleles based on SNPs at codons 127 and 172 in exon 5.

Allele	GH codon	
	127	172
A	CTG	ACG
B	GTG	ACG
C	GTG	ATG
D	CTG	ATG

We have provided evidence for the second scenario in Hosa. By studying the MHC, it was clear ancestral haplotypes were generated before the migration and expansion [17,18]. Thus we used the expression "conservation of polymorphism" to describe the freezing of ancestral haplotypes which occur within Polymorphic Frozen Blocks. It is true that there has been shuffling of haplotypes and some recombination, but very little change within the frozen sequences [19]. Some unstable sequences such as microsatellites have mutated and have generated some new SNPs but there is no evidence that such change is important in selection, fitness or any other functional sense. Rather, we argue that the polymorphism generated at the time of speciation events is the major determinant of fitness and that the survival and expansion of Hosa was more related to an ability to conserve polymorphism actively. Further evidence has been added by showing that the situation is similar in other Polymorphic Frozen Blocks such as the RCA region in Hosa [5] and the MHC of the domestic dog [11].

Here, we provide new data in Bota which supports the conserved model but might suggest that the two extreme models are not entirely exclusive. We show that there are conserved ancestral haplotypes at the MRIP region on C19. Clearly these preceded the migration and expansion of domestic cattle since they have been maintained from an earlier ancestral pool. On the other hand, we cannot exclude some recent mutation.

An interesting feature of Bota is the degree of shuffling through recombination. This could be partly explained by the fact that the region examined is some 14 Mb which is an order of magnitude greater than the Hosa MHC where haplotypes tend to fragment over hundreds of kilobases. It has been suggested, on statistical grounds, that recombination distances are greater in cattle [20] which could explain the present finding of haplotypes extending from SREBF1 to GH. Further studies of observed haplotypes at this and other regions are required before such conclusions are accepted.

Irrespective of the distances involved, recombinant shuffling has the potential to generate new haplotypes from existing sequences. We postulate that this is a major process leading to diversification during bottlenecks [4] and that the selection of certain recombinant haplotypes can assist in establishing new breeds for commercial purposes. The C19 MRIP region is clearly important functionally and may have been directly relevant to the intense mating selection based on phenotypes preferred by breeders. Impressive associations between muscle characteristics and alleles at SREBF1 and GH have been reported in Wagyu. There is no doubt that breeders pay handsomely for tests relating to these loci. It is less clear how these relate to functional phenotypes probably because much information is commercial in confidence and subject to patents.

We postulate that the MRIP region contains at least 2 closely linked polymorphic frozen blocks designated S and G which are contiguous in the same way that the Hosa MHC and RCA regions contain related blocks. Within these 2 blocks, there must be at least several loci affecting muscle structure and function. Although there must be other genomic regions which contribute, a simple hypothesis argues that Wagyu have MRIP haplotypes which result in impaired strength and

growth while favoring fat deposition within muscle. The present data show that SREBF1 and GH loci per se cannot explain this phenotype.

Recently, using GWA, Uemoto et al. 2010 have suggested that some 30 Mb of the C19 is relevant to low melting point fatty acid composition (oleic acid, C18:1) [21]. Interestingly, the 14 Mb around MRIP including SREBF1 and GH is included within the 30 Mb. GWA has major limitations [2,7]. The haplotyping described here will allow selective breeding for these traits.

The data suggest 3 stages in the evolution of cattle. An ancient progenitor population was characterized by several robust ancestral haplotypes which have survived. These same haplotypes have also given rise to more or less breed-specific haplotypes, largely by recombinant shuffling. Some alleles and haplotypes have been lost in some breeds. There is little evidence of recent mutation in the sense of the creation of new alleles with possible exceptions being NT5M 22 and GH C. Finally, breed specific haplotypes have been mixed by natural or intentional cross breeding again with shuffling.

4. Materials and methods

4.1. Cattle breeds

Pure breeds of European Simmental, Japanese Wagyu and British Angus cattle grown in the South Western region of WA at several locations: Melaleuka Stud and Roebrooks Farms (Simmental, Wagyu); and Blackrock and Esslemont (Aberdeen Angus). The Simmental herd was established in 1972. The Wagyu were first introduced from semen and embryos imported from Japan and then live Wagyu purchased from a local herd north of Perth.

The Zebu and Brahman herds were local.

4.2. DNA samples

Genomic DNA was extracted from whole blood collected into acid citrate dextrose using the standard salting out method. A total of 298 cattle of different breeds were tested including some three generation families to allow investigation of segregation.

4.3. Primer design

Sequences were extracted from the Bovine database (Oct 2007 (Baylor 4.0/bosTau4)) UCSC Genome Bioinformatics (http://genome.ucsc.edu/) and analyzed by algorithms which have been found to be helpful in demonstrating polymorphism in the absence of microsatellites, duplications and obvious indels [5]. The sequences were chosen because of well-defined polymorphisms with high heritability. The primers used are listed in Table 1.

4.4. PCR methods

All PCR reactions consisted of differing final concentrations of the following components: Qiagen Fast Cycling Master Mix 2×, Qiagen Q Solution 5×, genomic DNA, forward and reverse primers, and Qiagen double-distilled water in a total volume of 20 µl. All thermocycling was performed on a Corbett PCR machine (Fisher Biotec, Australia).

A multiplex reaction was utilized for SREBF1, TCAP and NT5M. Forward and reverse primers for each were used at final concentrations of 0.5, 0.5, and 0.75 µM, respectively. Sixty nanograms of genomic DNA was used per reaction and Qiagen Master Mix and Q solutions were diluted to 1×. Cycling parameters were as follows: 95 °C for 5 min; 35 cycles of 96 °C for 10 s, 60 °C for 30 s, and 68 °C for 45 s; followed by a final extension at 68 °C for 5 min then held at 15 °C.

Two amplifications were necessary for GH that utilized a common forward primer and four unique reverse primers with 80 ng of genomic DNA per reaction, and Qiagen Master Mix and Q solutions diluted to 1×. The first amplification was performed with the GH forward primer and GH reverse primers 1 and 3 and the second amplification incorporated the GH forward primer and GH reverse primers 2 and 4. The GH forward primer and reverse primers were used at 1 and 0.5 µM final concentrations respectively. Cycling parameters were the same as detailed for the multiplex reaction described above. Assignment of alleles is as shown in Table 6b.

Table 6b
Band patterns for particular haplotype combinations. Note AC and BD genotypes cannot be distinguished.

Genotype	127C	127G	172C	172T
AA	+	−	+	−
AB	+	+	+	−
AC	+	+	+	+
BD	+	+	+	+
AD	+	−	+	+
DD	+	−	−	+
CD	+	+	−	+
BC	−	+	+	+
BB	−	+	+	−
CC	−	+	−	+

To amplify the MRIP locus, forward and reverse primers at a final concentration of 0.5 µM were used along with 60 ng of DNA. Qiagen Fast Cycling Master Mix was diluted to 1×; however, the Qiagen Q Solution was used at 1.75×. A touchdown cycling program produced optimal amplification at this locus under the following conditions: 5 min at 95 °C; then 35 cycles with denaturation at 96 °C for 10 s, an annealing step lasting 30 s, then extension at 68 °C for 30 s. The annealing temperature for the first 5 cycles was 64 °C, then 5 cycles at 62 °C, 5 cycles at 60 °C, then 5 touchdown cycles from 59 to 55 °C, followed by 15 cycles at 55 °C. This was followed by a final extension at 68 °C for 5 min then held at 15 °C.

All PCR products were diluted 1:1 with Qiagen QX DNA dilution buffer before loading onto the QIAxcel high-resolution cartridge for separation and quantification using the QIAxcel system. Fresh QX (15/500 bp or 15 bp/1 kb) alignment markers were used in each sample. QX separation buffer and QX wash buffer were purchased from Qiagen Pty Ltd. Data were analyzed with Qiagen software.

Acknowledgments

We are grateful to Vivian Zhang, Rongchang Yang, Oscar Kalai, Patrick Carnegie and Brent Stewart for their suggestions and assistance. This is publication 0805 of the C Y O'Connor Foundation.

References

[1] A.L. Van Eenennaam, J. Li, R.M. Thallman, R.L. Quaas, M.E. Dikeman, C.A. Gill, D.E. Franke, M.G. Thomas, Validation of commercial DNA tests for quantitative beef quality traits, J. Anim. Sci. 85 (2007) 891–900.
[2] W.P. Smith, Q. Vu, S.S. Li, J.A. Hansen, L.P. Zhao, D.E. Geraghty, Toward understanding MHC disease associations: partial resequencing of 46 distinct HLA haplotypes, Genomics 87 (2006) 561–571.
[3] T. Shiina, H. Inoko, J.K. Kulski, An update of the HLA genomic region, locus information and disease associations: 2004, Tissue Antigens 64 (2004) 631–649.
[4] R. Dawkins, C. Leelayuwat, S. Gaudieri, G. Tay, J. Hui, S. Cattley, P. Martinez, J. Kulski, Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease, Immunol. Rev. 167 (1999) 275–304.
[5] J.F. Williamson, C.A. McLure, P.N. Baird, D. Male, J. Millman, B. Lawley, M.L. Ashdown, P.J. Keating, R.L. Dawkins, Novel sequence elements define ancestral haplotypes of the region encompassing complement factor H, Hum. Immunol. 69 (2008) 207–219.
[6] E.J. Yunis, C.E. Larsen, M. Fernandez-Vina, Z.L. Awdeh, T. Romero, J.A. Hansen, C.A. Alper, Inheritable variable sizes of DNA streches in the human MHC: conserved extended haplotypes and their fragments or blocks, Tissue Antigens 62 (2003) 1–20.
[7] C.A. Alper, C.E. Larsen, D.P. Dubey, Z.L. Awdeh, D.A. Fici, E.J. Yunis, The haplotype structure of the human major histocompatibility complex, Hum. Immunol. 67 (2006) 73–84.

[8] S. Hoashi, N. Ashida, H. Ohsaki, T. Utsugi, S. Sasazaki, M. Taniguchi, K. Oyama, F. Mukai, H. Mannen, Genotype of bovine sterol regulatory element binding protein-1 (SREBP-1) is associated with fatty acid composition in Japanese Black cattle, Mamm. Genome 18 (2007) 880–886.
[9] S. Tsuji, Method for determining gene relating to favorable beef taste and texture, E.P.A. EP1845159, 2007.
[10] Z. Jiang, J.J. Michal, D.J. Tobey, T.F. Daniels, D.C. Rule, M.D. Macneil, Significant associations of stearoyl-CoA desaturase (SCD1) gene with fat deposition and composition in skeletal muscle, Int. J. Biol. Sci. 4 (2008) 345–351.
[11] C.A. McLure, P.W. Kesners, S. Lester, D. Male, C. Amadou, J.R. Dawkins, B.J. Stewart, J.F. Williamson, R.L. Dawkins, Haplotyping of the canine MHC without the need for DLA typing, Int. J. Immunogenet. 32 (2005) 407–411.
[12] N. Longman-Jacobsen, J.F. Williamson, R.L. Dawkins, S. Gaudieri, In polymorphic genomic regions indels cluster with nucleotide polymorphism: quantum genomics, Gene 312 (2003) 257–261.
[13] S. Gaudieri, R.L. Dawkins, K. Habara, J.K. Kulski, T. Gojobori, SNP profile within the Human Major Histocompatibility Complex reveals an extreme and interrupted level of nucleotide diversity, Genome Res. 10 (2000) 1579–1586.
[14] K. Kitcharoen, C.S. Witt, A.V. Romphruk, F.T. Christiansen, C. Leelayuwat, MICA, MICB, and MHC beta block matching in bone marrow transplantation: relevance to transplantation outcome, Hum. Immunol. 67 (2006) 238–246.
[15] M.A. Degli-Esposti, L.J. Abraham, V. McCann, T. Spies, F.T. Christiansen, R.L. Dawkins, Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM, Immunogenetics 36 (1992) 345–356.
[16] M.A. Degli-Esposti, A. Andreas, F.T. Christiansen, B. Schalke, E. Albert, R.L. Dawkins, An approach to the localization of the susceptibility genes for generalized myasthenia gravis by mapping recombinant ancestral haplotypes, Immunogenetics 35 (1992) 355–364.
[17] S. Gaudieri, N. Longman-Jacobsen, G.K. Tay, R.L. Dawkins, Sequence analysis of the MHC class I region reveals the basis of the genomic matching technique, Hum. Immunol. 62 (2001) 279–285.
[18] M.A. Degli-Esposti, A.L. Leaver, F.T. Christiansen, C.S. Witt, L.J. Abraham, R.L. Dawkins, Ancestral haplotypes: conserved population MHC haplotypes, Hum. Immunol. 34 (1992) 242–252.
[19] S. Gaudieri, C. Leelayuwat, G.K. Tay, D.C. Townend, R.L. Dawkins, The major histocompatibility complex (MHC) contains conserved polymorphic genomic sequences that are shuffled by recombination to form ethnic-specific haplotypes, J. Mol. Evol. 45 (1997) 17–23.
[20] M.E. Goddard, B.J. Hayes, Mapping genes for complex traits in domestic animals and their use in breeding programmes, Nat. Rev. Genet. 10 (2009) 381–391.
[21] Y. Uemoto, T. Abe, N. Tameoka, H. Hasebe, K. Inoue, H. Nakajima, N. Shoji, M. Kobayashi, E. Kobayashi, Whole-genome association study for fatty acid composition of oleic acid in Japanese Black cattle, Anim. Genet. 42 (2011) 141–148.
[22] K. Chikuni, Y. Fukumoto, R. Tanabe, S. Muroya, S. Ozawa, A simple method for genotyping the bovine growth hormone gene, Anim. Genet. 28 (1997) 230–232.